FUNDAMENTAL IMMUNOLOGY

SECOND EDITION

Editor

WILLIAM E. PAUL, M.D.

Laboratory of Immunology
National Institute of Allergy and
Infectious Diseases
National Institutes of Health
Bethesda, Maryland

Raven Press New York

Raven Press, 1185 Avenue of the Americas, New York, New York 10036

Made in the United States of America

Library of Congress Cataloging-in-Publication Data

Fundamental immunology.

Includes bibliographical references and index.
1. Immunology. I. Paul, William E. [DNLM:
1. Immunity. QW 504 F9804]
QR181.F84 1989 616.07'9 85-43216
ISBN 0-88167-491-5

The material contained in this volume was submitted as previously unpublished material, except in the instances in which some of the illustrative material was derived.

Great care has been taken to maintain the accuracy of the information contained in this volume. However, neither Raven Press nor the editor can be held responsible for errors or for any consequences arising from the use of the information contained herein.

9 8 7 6 5 4 3 2 1

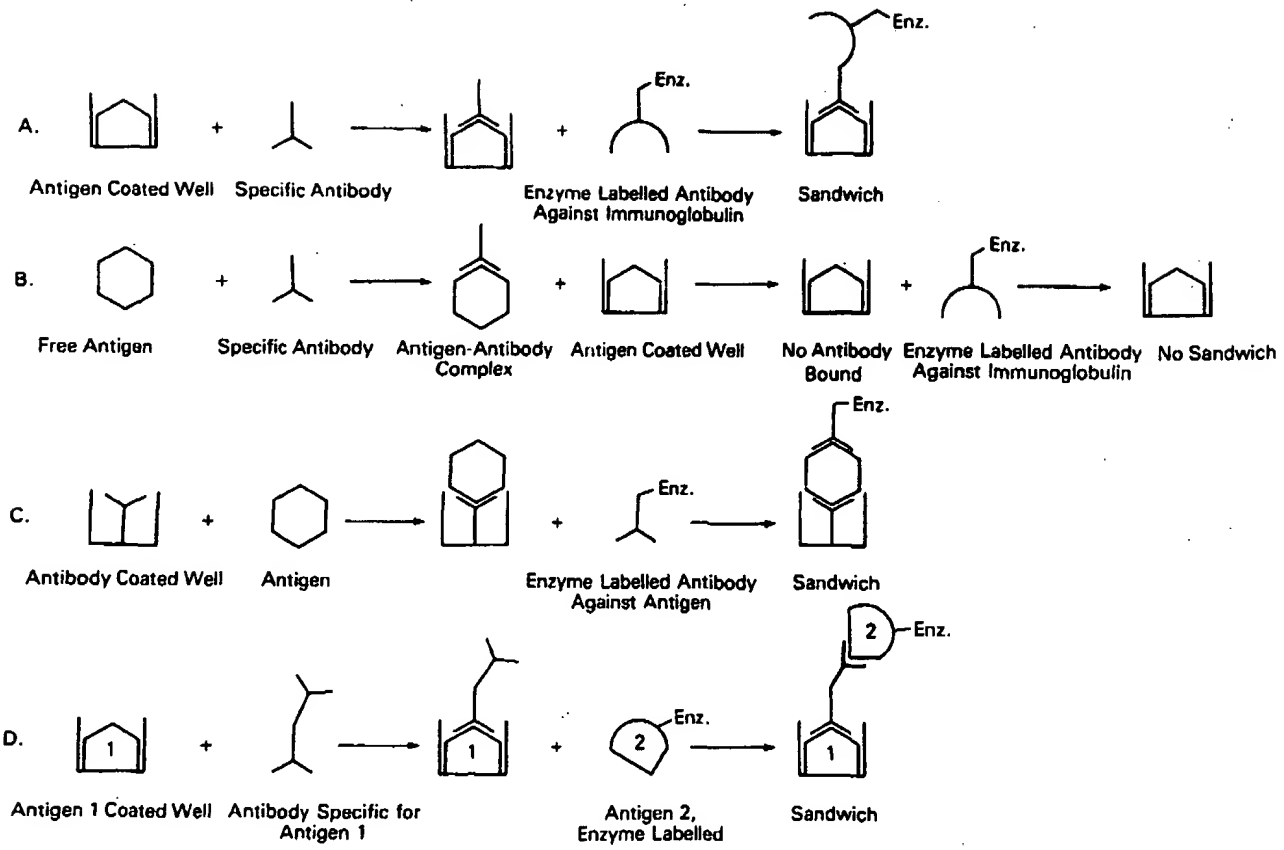

FIG. 9. Four strategies for the detection of specific antibody–antigen reactions using the ELISA technique. **A:** Direct binding. **B:** Hapten inhibition. **C:** Antigen sandwich. **D:** Antibody sandwich.

rations of idiotype, or for detection of anti-anti-idiotype (see Chapters 9 and 23). Other arrangements of antibody and antigen are also possible. Extra layers of detecting reagents can amplify sensitivity but also tend to raise the background and introduce variability.

An example of the first method described above is the detection of human antibodies to influenza virus (52) (Fig. 10). Alternate columns were coated with influenza virus or bovine albumin. Serum was added at 1/10 dilution to the top two wells of each box and serially diluted in four-fold steps from top to bottom. The last colored well indicates the titer, whereas the absence of color in the albumin-coated wells indicates the specificity. A second use of this method is for screening culture supernatants in the production of hybridoma antibodies. The sensitivity and speed of the ELISA method make it possible to screen large numbers of wells for the production of specific antibody. Clones selected by this method tend to have high antigen affinities, perhaps due to dissociation of low-affinity antibodies during the wash steps.

An important caution when using native protein antigens to coat solid-phase surfaces (Fig. 9A) is that binding to a surface can alter the conformation of the protein. For instance, using conformation specific monoclonal antibodies to myoglobin, Darst et al. (53) found that binding of myoglobin to a surface altered the apparent affinity of some antibodies more than others. This problem may be avoided by using the method of Fig. 9B.

SPECIFICITY AND CROSS-REACTIVITY

The specificity of an antibody or antiserum is defined by its ability to discriminate between the antigen against which it was made (called the homologous antigen, or immunogen) and any other antigen one might test. In practice, one cannot test the whole universe of antigens, but only selected antigens. In this sense, specificity can only be defined experimentally within that set of antigens one chooses to compare. Karush (28) has defined a related term, selectivity, as the ability of an antibody to discriminate, in an all-or-none fashion, between two related ligands. Thus selectivity depends not only on the relative affinity of the antibody for the two ligands but on the experimental lower limit for detection of reactivity. For instance, an anti-carbohydrate antibody with an affinity of $10^5\ M^{-1}$ for the immunogen may appear to be highly selective, since reaction with a related carbohydrate with a 100-fold lower affinity, $10^3\ M^{-1}$, may be undetectable. On the other hand, an antibody with an affinity of $10^9\ M^{-1}$ for the homologous ligand may appear to be less selective because any reaction with a related ligand with a 100-fold lower affinity would still be quite easily detectable.

Conversely, cross-reactivity is defined as the ability to react with related ligands other than the immunogen.

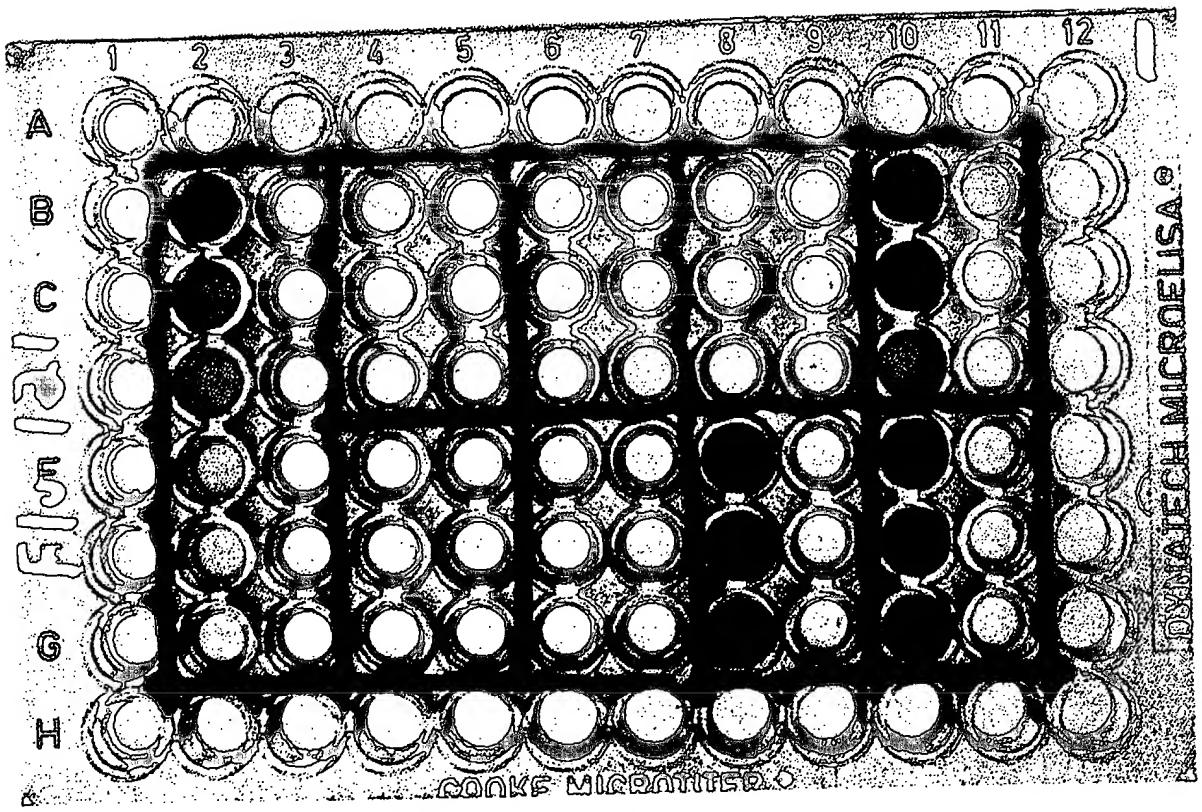

FIG. 10. Typical ELISA plate, using the direct binding method. Wells in alternate columns were coated with influenza virus or bovine serum albumin. Samples to be tested for anti-influenza antibodies were added to the top wells of each box and serially diluted in fourfold steps from top to bottom. A control titration of a known standard antiserum was performed in columns 2 and 3. Enzyme-labeled goat anti-human Ig was used as detecting reagent. (Photograph courtesy of Dr. R. Yarchoan and Dr. D. L. Nelson.)

More usually, this is examined from the point of view of the ligand. Thus one might say that antigen Y cross-reacts with antigen X because it binds to anti-X antibodies. Note that in this sense, it is the two antigens that are cross-reactive, not the antibody. However, the cross-reactivity of two antigens, X and Y, can be defined only with respect to a particular antibody or antiserum. For instance, a different group of anti-X antibodies may not react at all with Y, so that with respect to these antibodies, Y would not be cross-reactive with X. One can also use the term in a different sense, saying that some anti-X antibodies cross-react with antigen Y.

In most cases, cross-reactive ligands have lower affinity than the immunogen for a particular antibody. However, exceptions can occur, in which a cross-reactive antigen binds with a higher affinity than the homologous antigen itself. This phenomenon is called heteroclicity, and the antigen that has a higher affinity for the antibody than does the immunogen is said to be heteroclitic. Antibodies that manifest this behavior are also described as heteroclitic antibodies. A good example is the case of antibodies raised in C57BL/10 mice against the hapten nitrophenyl acetyl (NP). These antibodies have been shown by Mäkelä and Karjalainen (54) to bind with higher affinity to the cross-reactive hapten, nitroiodophenyl acetyl (NIP), than to the immunogen itself.

In many practical situations, cross-reactivity is detected by methods such as precipitin, especially precipitation in agar (the Ouchterlony test), or hemagglutination (see below for descriptions of both of these) or similar methods which have in common the fact that they do not distinguish well between differences in affinity and differences in concentration. This practical aspect, coupled with the heterogeneity of immune antisera, has led to ambiguities in the usage of the terms "cross-reactivity" and "specificity." With the advent of RIA techniques, this ambiguity in the terminology, as well as in the interpretation of data, has become apparent.

For these reasons, Berzkofsky and Schechter (55) have recently defined two forms of cross-reactivity, and, correspondingly, two forms of specificity. These two forms of cross-reactivity are illustrated by the two prototype competition RIA curves in Fig. 11. In reality, most antisera display both phenomena simultaneously.

Type 1 cross-reactivity, or true cross-reactivity, is defined as the ability of two ligands to react with the same site on the same antibody molecule, possibly with different affinities. For example, the related haptens dinitro-

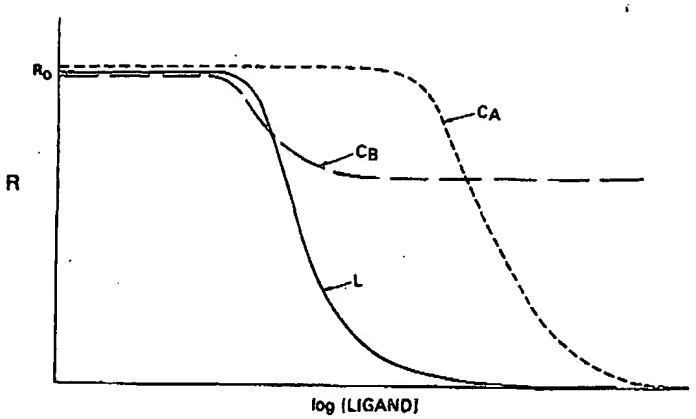

FIG. 11. Schematic RIA binding curves for homologous ligand L and cross-reacting ligands. Cross-reacting ligand C_A manifests Type I or true cross-reactivity demonstrated by complete inhibition of tracer ligand binding, and a lower affinity. Ligand C_B displays Type II cross-reactivity or determinant sharing, as recognized from the plateau at less than 100% inhibition, but not necessarily a lower affinity. The ordinate R is the ratio of bound/free radiolabeled tracer ligand, and R_0 is the limit of R as the concentration of all ligands, including tracer, approaches zero. (From Berzofsky and Schechter, ref. 55.)

phenyl (DNP) and trinitrophenyl (TNP) may react with different affinity for antibodies raised to dinitrophenyl hapten. In protein antigens, such differences could occur with small changes in primary sequence (e.g., the conservative substitution of threonine for serine), or with changes in conformation, such as the cleavage of the protein into fragments (Fig. 12). If a peptide fragment contained all the contact residues in an antigenic determinant (i.e., those which contact the antibody-combining site), it might cross-react with the native determinant for antibodies against the native form, but with lower affinity because the peptide would not retain the native conformation (see Chapter 8). This type of affinity difference is illustrated by competitor C_A in Fig. 11, in which complete displacement of tracer can be achieved at high enough concentrations of C_A, but higher concentrations of C_A than of the homologous ligand, L, are required to produce any given degree of inhibition.

A separate issue from affinity differences is the issue of whether the cross-reactive ligand reacts with all or only a subpopulation of the antibodies in a heterogeneous serum. This second type of cross-reactivity, which we call Type 2 cross-reactivity or shared reactivity, therefore can occur only when the antibody population is heterogeneous, as in most conventional antisera. In this case, the affinity of the cross-reactive ligand may be greater than, less than, or equal to that of the homologous ligand for those antibodies with which it interacts. Therefore, the competition curve is not necessarily displaced to the right, but the inhibition will reach a plateau at less than complete inhibition, as illustrated by competitor C_B in Fig. 11. As an example, let us consider the case of a protein with determinants X and Y, and an antiserum against this protein containing both anti-X and anti-Y antibodies. Then a mutant protein in which determinant Y was so altered as to be unrecognizable by anti-Y, but determinant X was intact, would manifest Type 2 cross-reactivity. It would compete with the wild-type protein only for anti-X antibodies (possibly even with equal affinity), but not for anti-Y antibodies.

Occasionally, even monoclonal antibodies may appear to display Type 2 cross-reactivity in situations in which secondary reactions are involved in the measurement of the antigen–antibody reaction. For example, Sharon et al. (60) and Cisar et al. (61) observed plateau values at less than 100% binding of a homogeneous myeloma or hybridoma antibody reactive with dextrans. In this case, the assay used was quantitative precipitin, in which differential solubility of different complexes could account for such a plateau. If one could directly observe the antigen–antibody interaction in solution, without the need for any secondary reaction that might be incomplete, the reaction of a homogeneous antibody with its homogeneous antigen theoretically cannot reach a plateau at less than 100% reaction or inhibition. Therefore the existence of secondary competing reactions should be considered when such plateaus are observed.

Of course, both types of cross-reactivity could occur simultaneously. A classic example would be the peptide fragment discussed in the case of Type I cross-reactivity above. Suppose the fragment contained the residues of determinant X, albeit not in the native conformation, but did not contain the residues of a second determinant, Y, which was also expressed on the native protein. If the antiserum to the native protein consisted of anti-X and anti-Y, the peptide would compete only for anti-X antibodies (Type 2 cross-reactivity) but would have a lower affinity than the native protein even for these antibodies. Thus the competition curve would be shifted to the right and would plateau before reaching complete inhibition.[8]

In the case of a homogeneous (e.g., monoclonal) antibody in which only Type 1 or true cross-reactivity can occur, one can quantitate the differences in affinity for different cross-reactive ligands by a method analogous to the B/F versus F method described above. Suppose that ligands X and Y cross-react with homologous ligand L for a monoclonal antibody. If one plots the bound/free ($B/F = R$) ratio for radiolabeled tracer ligand L as a function of the log of the concentration of competitors X and Y, one obtains two parallel competition curves (Fig. 13), under the appropriate conditions (below). The first con-

[8] An ambiguous case could occur experimentally in which the distinction between the two types of cross-reactivity would be blurred. For example, in the case of antibodies that all react with determinant X but have a very wide range of affinities for X, some such antibodies may have such a low affinity for cross-reactive determinant X′ that they would appear not to bind X′ at all. Then a competition curve using X′ might appear to reach a plateau at incomplete inhibition even though all the antibodies were specific for X, and the only difference between X and X′ was affinity.

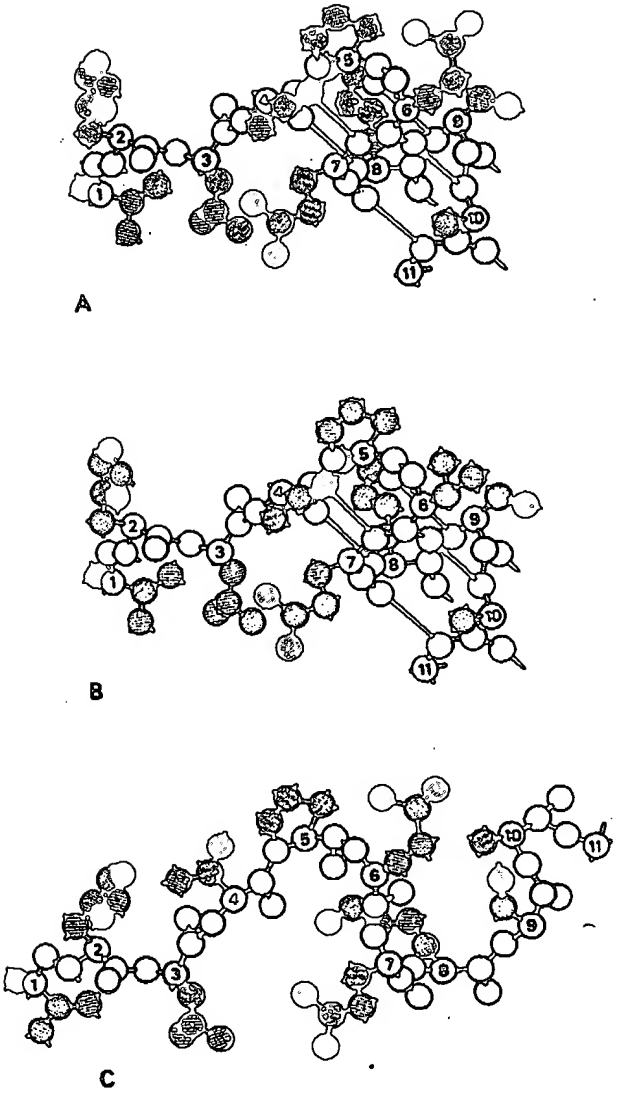

FIG. 12. An artist's drawing of the amino terminal region of the β chain of hemoglobin. A: The first 11 residues of the β^A chain. B: The comparable regions of the β^S chain. The substitution of valine for the normal glutamic acid at position 6 makes a distinct antigenic determinant to which a subpopulation of antibodies may be isolated (56,57). C: A schematic diagram of the sequence in A, unfolded as occurs when the protein is denatured. This region may be cleaved from the protein, or the peptide synthesized (58), resulting in changed antigenic reactivity. An antiserum prepared to hemoglobin (or the β chain thereof) might exhibit cross-reactivity with the structures shown in B and C but the molecular mechanisms would be different. Polypeptide backbone atoms are in *white*; in the side chains, oxygen atoms are *hatched*, nitrogen atoms are *black*, and carbon atoms are *lightly stippled*. (Adapted from Dean and Schechter, ref. 59, and Berzofsky and Schechter, ref. 55.)

dition is that the concentration of free tracer be less than $1/K_L$, the affinity for tracer. In this case, it can be shown (55) that

$$K_X \simeq \frac{1}{[X]_{free}} \tag{41}$$

at the midpoint where $R = R_0/2$, where K_X is the affinity for X. This is analogous to Equation (21) for the case in which unlabeled homologous ligand is the competitor. Also, in analogy with Equation (23), it can be shown that if the total concentration of competitor, $[X]_t$, is used instead of the free concentration, $[X]_{free}$, an error term will arise, giving

$$[X]_t \text{ (at } R = R_0/2) = \frac{1}{K_X} + \frac{[S]_t}{2} \tag{42}$$

Thus, with competitor on a linear scale, the difference in midpoint for competitors X and Y will correspond to the difference $1/K_X - 1/K_Y$ regardless of whether free or total competitor is plotted, but the ratio of midpoint concentrations will equal K_X/K_Y only if the free concentrations are used. This last point is important if one plots the log of competitor concentration, as is usually done, since the horizontal displacement between the two curves on a log scale corresponds to the ratio [X]/[Y], not the difference (55).

If a second condition also holds, namely, that the concentration of bound tracer is small compared to the antibody site concentration $[S]_t$, then the slopes (on a linear scale) of the curves at their respective midpoints (where $R = R_0/2$) will be proportional to the affinity for that competitor, K_X or K_Y (55). (Both conditions can be met by keeping tracer L small relative to both K_L and $[S]_t$.) When $[X]_{free}$ and $[Y]_{free}$ are plotted on a log scale, the slopes will appear to be equal (i.e., the curves will appear parallel), since a parallel line shifted m-fold to the right on a log scale will actually be $1/m$ as steep, at any point, in terms of the antilog as abscissa.

When the antibodies are heterogeneous in affinity, the curves will be broadened and in general will not be parallel. When heterogeneity of specificity is present, and Type 2 cross-reactivity occurs, it should be pointed out that the fractional inhibition achieved at the plateau in a B/F versus free competitor plot will not be proportional to the fraction of antibodies reacting with that competitor but will be proportional to a weighted fraction, where the antibody concentrations are weighted by their affinity for the tracer (55).

These two types of cross-reactivity lead naturally to two definitions of specificity (55). The overall specificity of a heterogeneous antiserum is a composite of both of these facets of specificity. Type 1 specificity is based on the relative affinities of the antibody for the homologous ligand and any cross-reactive ligands. If the affinity is much higher for the homologous ligand than for any cross-reactive ligand tested, then the antibody is said to be highly specific for the homologous ligand; that is, it discriminates very well between this ligand and the others. If the affinity for cross-reactive ligands is below the threshold for detection in an experimental situation, then

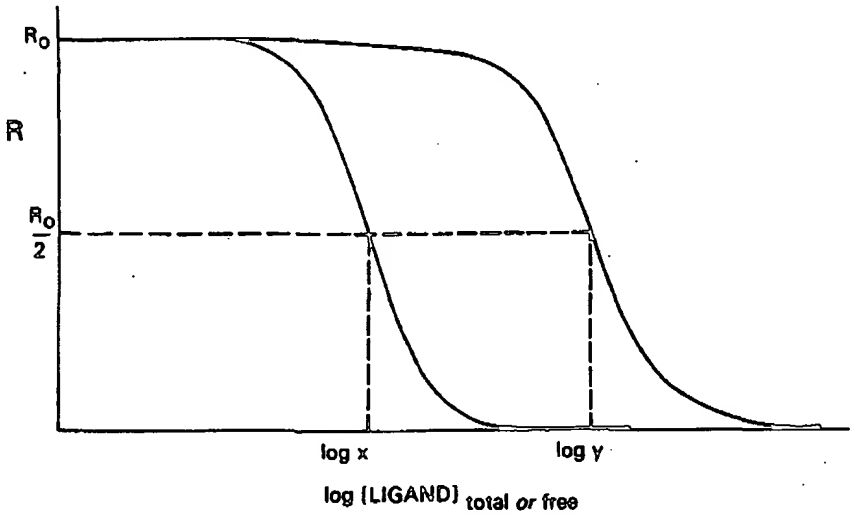

FIG. 13. Schematic RIA binding curves showing the effect of affinity on the midpoint and the slope at the midpoint, and the value of using free [ligand] rather than total [ligand]. Ordinate R is the ratio of bound/free radiolabeled tracer ligand, and R_0 the limit of R as all ligand concentrations approach zero. If x and y are the concentrations of ligands X and Y which reduce R to exactly $R_0/2$, then if the abscissa is total ligand concentration, $x = 1/K_X + [S]_t/2$ and $y = 1/K_Y + [S]_t/2$, where $[S]_t$ is the concentration of antibody binding sites and K_X and K_Y the affinities of the antibody for the respective ligands. However, if the abscissa is free ligand concentration, $x = 1/K_X$ and $y = 1/K_Y$ so that the ratio x/y (or the difference log x − log y on a log plot) corresponds to the ratio of affinities K_Y/K_X. Note that the slopes at the midpoints are the same on a log scale, but that for Y would be only K_Y/K_X that for X on a linear scale. (From Berzofsky and Schechter, ref. 55, with permission.)

Type 1 specificity gives rise to selectivity as was discussed above (cf. ref. 28). The specificity can even be quantitated in terms of the ratio of affinities for the homologous ligand and a cross-reactive ligand (cf. ref. 62). It is this Type 1 specificity that most immunochemists would call true specificity, just as we have called Type 1 cross-reactivity true cross-reactivity.

The common use of the term "cross-reactivity" to include Type 2 or partial reactivity leads to a second definition of specificity which applies only to heterogeneous populations of antibodies such as antisera. We call this Type 2 specificity. If all the antibodies in the mixture react with the immunogen, but only a small proportion react with any single cross-reactive antigen, then the antiserum would be said to be relatively specific for the immunogen. Note that it does not matter whether the affinity of a subpopulation which reacts with a cross-reactive antigen is high or low (Type 1 cross-reactivity). As long as that subpopulation is a small fraction of the antibodies, the mixture is specific. Thus Type 2 specificity depends on the relative concentrations of antibodies in the heterogeneous antiserum, not just on their affinities. Also note that one can use these relative concentrations of antibody subpopulations to compare the specificity of a single antiserum for two cross-reactive ligands. However, it would not be meaningful to compare the specificity of two different antisera for the same ligand by comparing the fraction of antibodies in each serum which reacted with that ligand. Although Type 2 specificity may appear to some a less classic concept of specificity than Type 1, it is Type 2 specificity that one primarily measures in such assays as the Ouchterlony double immunodiffusion test, and it carries equal weight with Type 1 specificity in such assays as hemagglutination, discussed below. Type 2 specificity also leads naturally to the concept of "multispecificity" described below.

Multispecificity

The theory of multispecificity, introduced and analyzed by Talmadge (63) and Inman (64,65) and discussed on a structural level by Richards et al. (66), suggests a mechanism by which the great diversity and specificity of antisera can be explained without the need for a correspondingly large repertoire of antibody structures (or structural genes). The idea is that each antibody may actually bind, with high affinity, a wide variety of quite diverse antigens. When one immunizes with immunogen A, one selects for many distinct antibodies which have in common only that they all react with A. In fact, each antibody may react with other compounds, but if fewer than 1% of the antibodies bind B, and fewer than 1% bind C, and so on, then by Type 2 specificity, the whole antiserum will appear to be highly specific for A. Note that the subpopulation which binds B may react with an affinity for B as high as or higher than that for A, so that the population would not be Type 1 specific for A. This same population would presumably be selected if one immunized with B, as well as with perhaps hundreds of other immunogens with which these antibodies react. The net result would be that